PRTN-283

22

ACTIVITY SIGNALLING AND IMPROVED HOP ACKNOWLEDGMENTS IN PACKET RADIO SYSTEMS

P. Spilling, SRI International
F. Tobagi, Stanford University
January 1980

I INTRODUCTION

The current Packet Radio protocol utilizes the carrier sense technique to detect channel activity, as a decision criterium whether to send a packet or wait. This method works well in a fully connected network where all participating Packet Radio units can hear each other, but fails in multi-hop networks, where the traffic sources may be located more than one hop away, and therefore unable to detect each others activity. This is called the Hidden Terminal problem. The result is a finite probability for collisions between packets transmitted simultaneously from more than one Packet Radio unit, with the result that throughput is reduced and delay is increased.

It is assumed that time-overlapping receptions of packets are garbled and received in error. This is not always true. The Packet Radios employ a spread spectrum technique, which means that one receiver may lock on to a particular transmission and neglect others, if that transmission is first in time, or has a high received energy compared to the other contending transmissions. This implies that all but one transmission are unsuccessful, and will be retransmitted later. It is assumed that these unnecessary retransmissions are contributing to the channel noise, thus increasing the probability for errors in packets being received.

The Packet Radio protocol utilizes a hop-by-hop acknowledgment scheme. A hop transmission can either be acknowledged by the packet being forwarded over the next hop, called echo-ACK, or by a dedicated

ACK-packet, called active-ACK, when the hop transmission was the last one on the packet route. The Hidden Terminal problem has a substantial effect, at high traffic loads, on the efficiency of the acknowledgment scheme, resulting in unnecessary retransmissions and loading of the channel, and unnecessary prolonged allocation of scarce buffer resources in the Packet Radio units.

This note discusses the effects induced by the Hidden Terminal problem on throughput and delay, and on the acknowledgment scheme, and proposes a method to eliminate or reduce these effects. The method can be viewed as an extension of the carrier sense method, by means of signalling of ongoing transmission and reception activities. The note does not discuss ways to implement this scheme, but it is believed that there are several possibilities. These approaches are currently being explored. In addition, other methods to improve the acknowledgment efficiency are discussed.

The following section discusses the Hidden Terminal problem and its effects on packet transmissions and acknowledgments. Then the principle of the activity signalling is discussed showing how this can be used to reduce these effects. Several possible approaches to improve the acknowledgment protocol are outlined in the section thereafter. And finally follows the requirements for the proposed activity signalling and acknowledgment scheme with some concluding remarks. A list of reference articles is included at the end of the note.

II THE HIDDEN TERMINAL PROBLEM

Consider a hop-transmission, from S to D, somewhere on the packet route between its ultimate source and destination, as shown in Figure 1. The two circles, A and B, represent the range of the packet transmissions from S and D, respectively. All Packet Radio units inside the circle A, called the ensemble A, are neighbours of S since they can all hear S. The ensemble B are all neighbours of D. Assume now that S

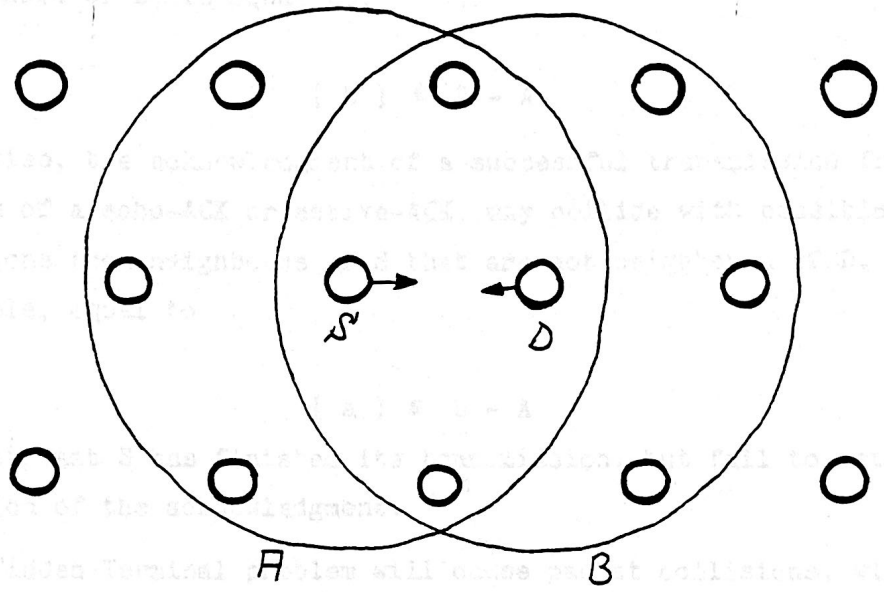

FIGURE 1 PART OF A PACKET RADIO NETWORK
RELEVANT TO THE HIDDEN TERMINAL PROBLEM

finds the channel empty, i.e., no carrier is detected, and transmits a packet destined for D. The ensemble A will then detect this and under carrier sense access rules will be inhibited from transmitting as long as S continues. There will of course be a vulnerable interval here, from the moment S detects that the channel is empty and starts the packet transmission until all its neighbours have detected the transmission from S. This period will be equal to the time it takes for a packet to propagate from S to its most distant neighbours. The latter neighbours can also have started packet transmissions less than one propagation time ahead of S, and not being detected by S, when it started its transmission. The total uncertainty interval is therefore equal to twice the propagation time from S to its most distant neighbours, and is called the diameter of A.

Since the packet from S is destined for D, packets transmitted from one or more of D's neighbours which are not neighbours of S (i.e., those

not detecting the packet transmission) will collide with the packet in question, making them all unsuccessful. This collection of neighbours, a sub-ensemble of B, is equal to:

$$\{ b \} \in B - A$$

Likewise, the acknowledgment of a successful transmission from S to D by means of a echo-ACK or active-ACK, may collide with possible packet transmissions from neighbours of S that are not neighbours of D. This sub-ensemble, equal to

$$\{ a \} \in B - A$$

will detect that S has finished its transmission, but fail to detect the transmission of the acknowledgment.

The Hidden Terminal problem will cause packet collisions, with a probability that depends upon the network configuration, the traffic pattern, and the traffic level. These collisions result in reduced throughput, increased delay, and increased load on the channel due to the extra retransmissions. The effects of the Hidden Terminal problem on packet transmissions and acknowledgments may be calculated analytically when certain assumptions are made about the distribution of the Packet Radio units and the traffic pattern. In the general case these effects must be obtained by means of simulations.

III ACTIVITY SIGNALLING

A. The Principle of Activity Signalling

In order for a transmission, as shown in Figure 1, to be successful, the packet must be successfully received by D, and its acknowledgment successfully received by S. This implies that none of D's neighbours may transmit simultaneously with S. Only the sub-ensemble containing the common elements of A and B, $A \cap B$, will be able to detect that S is active and be inhibited from transmitting. The rest of the ensemble B, B - A, will therefore be able to transmit. The purpose of activity signalling is to inform this sub-ensemble about the transmission from S to D; and when the transmission is over, to prevent any of S's neighbours from transmitting in an interval afterwards, so that the acknowledgment, in the form of echo-ACK or active-ACK, may be successful. The method should disturb the rest of the network as little as possible.

The principle of activity signalling, called Busy Tone, is illustrated in Figure 2, for an intermediate hop on the packet route, between source and destination. When S finds the channel empty, i.e., no carrier or Busy Tone is detected, it transmits the packet. At the end of the transmission, S turns on the Busy Tone for an interval T, equal to the expected time for D to process and forward the packet, or transmit an active-ACK. This inhibits transmissions of packets from any of its neighbours, after the packet transmission is finished and the channel becomes empty, and assures successful reception of the acknowledgment. The case in which D is delayed in responding to the received packet, is dealt with in the next section.

D, when hearing the packet from S, turns on its Busy Tone to prevent any of its neighbours from transmitting and colliding with the packet being received. The Busy Tone is turned off when the packet has been completely received.

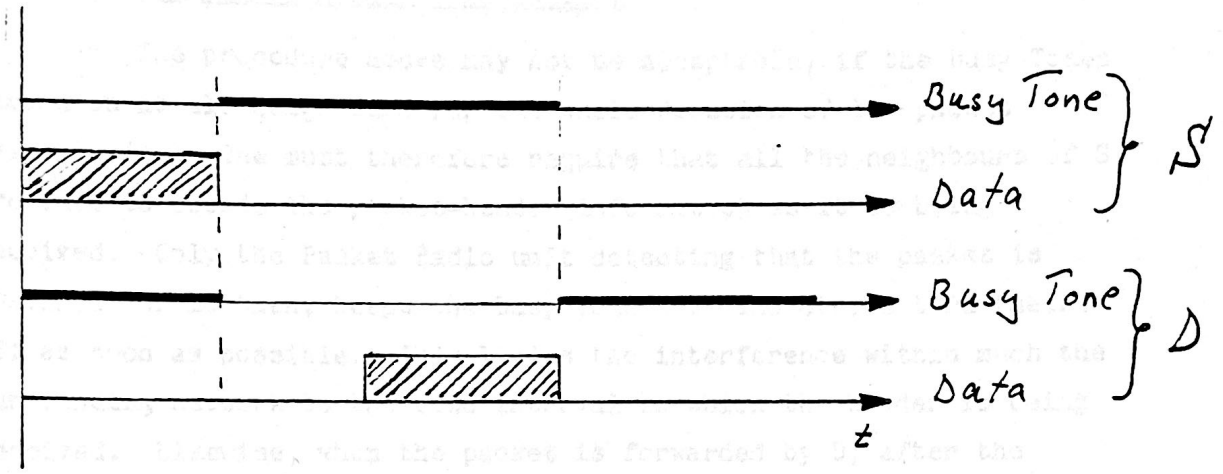

FIGURE 2 THE PRINCIPLE OF THE BUSY TONE ON A HOP TRANSMISSION

B. Discussions

There are different ways to reallize these intentions. They are discussed below.

1. Case 1 All neighbours of S turn on their busy tones immediately after detecting the transmisson from S.

The scheme requires that all Packet Radio units should turn on their Busy Tones immediately after detecting a packet transmission (on preamble detection for example). This implies that all neighbours of S, including D, turn on their Busy Tones and may thereby inhibit a larger portion, of the total network, than necessary.

2. Case 2 All neighbours of S turn on their busy tones immediately after detecting the transmission from S, and all but D turn them off again after reception and processing of the packet header information.

The procedure above may not be acceptable, if the Busy Tones remain on at all neighbours for the whole duration of the packet transmission. One must therefore require that all the neighbours of S are able to decode the packet-header information as it is being received. Only the Packet Radio unit detecting that the packet is destined for it then, keeps the Busy Tone on. The others turn theirs off as soon as possible. This limits the interference within much the surrounding network to the time interval in which the header is being received. Likewise, when the packet is forwarded by D, after the necessary processing time, all neighbours of D will turn on their Busy Tones for a time interval corresponding only to the transmission of the header information. The disturbance on the whole network may be substantial, but limited in time.

In the absence of a separate header checksum, header errors may cause that an unintended Packet Radio unit keeps its Busy Tone on until the whole packet has been received.

3. Case 3 D turns on the Busy Tone after reception and processing of the packet header information.

All neighbours of S first process the header of the packet being received, and only the destination D turns on its Busy Tone at that time. This inhibits just the necessary portion of the network, but has a vulnerable period approximately equal to the transmission time of the header portion of the packet.

As in the previous case, header errors may cause a wrong Packet Radio unit to turn on its Busy Tone, or that more units do so.

4. Case 4 No Busy Tone turned on by the source S.

It is not necessary for the source S to turn on its Busy Tone, during or after the packet transmission. All neighbours of S that are able to detect the packet being transmitted, would know that S is busy, and that they should wait a given interval after the packet transmission, before they start any transmission. This requires that they all continuously monitor the channel activity, i.e., carrier detect, and know when S is finished and the given interval has expired. If this is not possible, S needs to turn on its Busy Tone at the end of the packet transmission.

5. Case 5 Destination address information conveyed via the Busy Tone.

At first glance, this method may seem promising. But it makes the technique much more complicated, and requires that the source starts to transmit the Busy Tone a given interval prior to the actual packet transmission. This interval must be equal to the time needed to transmit the address information and for the destination to decode it, so that the destination turns on its Busy Tone simultaneously with the packet transmission from the source. Time-overlapping Busy Tones are assumed to be garbled, thus no information can be conveyed in that period. Since all the information needed for the destination, is contained in the packet, there is not much gained in coding the Busy Tone.

6. Comments

It would have been possible, in principle, for some neighbours of S, the members of the sub-ensemble A - B, to transmit a packet simultaneously with that from S, that was destined for a non-neighbour of S. Such a transmission would not interfere with the packet reception at D, and likewise the transmission from S would not interfere with its non-neighbours reception. In the case where the transmission from S is short, and the other long, there will be a certain probability that the

latter transmission may overlap the acknowledgment from D (echo-ACK or active-ACK) and make it unsuccessful. The strategy would therefore be to inhibit all transmission attempts during the packet transmission from S and the succeeding time interval thereafter.

The first method interferes with a larger portion of the network than necessary. Though limited in time, the disturbance may be substantial. The second method limits the interference, with the surrounding network, to the necessary portion, but has a larger vulnerable interval than the first one. Which of the two methods should be preferred, is a difficult question to answer, since the effects caused by either methods depend upon the network configuration, the traffic pattern, and traffic level.

IV IMPROVED HOP ACKNOWLEDGMENT

The Packet Radio units have limited buffer resources available for keeping packets or associated information. It is therefore essential that the acknowledgment method is as efficient as possible, so that buffer resources are not allocated longer than necessary. Prolonged allocations may eventually lead to lack of buffers, and will have a substantial effect on throughput and delay.

Acknowledgments that are lost or delayed too long, will in addition, result in unnecessary retransmissions, which will reduce throughput and channel utilization, and increase delays.

As mentioned previously, acknowledgments are performed, either via the header information in the packet forwarded over the next hop, or if the current hop is the last one, by transmitting an active-ACK.

In the current Packet Radio protocol, CAP 5, the packet header is processed after the whole packet has been received and verified for correctness. Measurements show that this introduces considerable delays. This is illustrated in Figure 3, for a full-size packet (approximately 2100 bits) being forwarded over an intermediate hop.

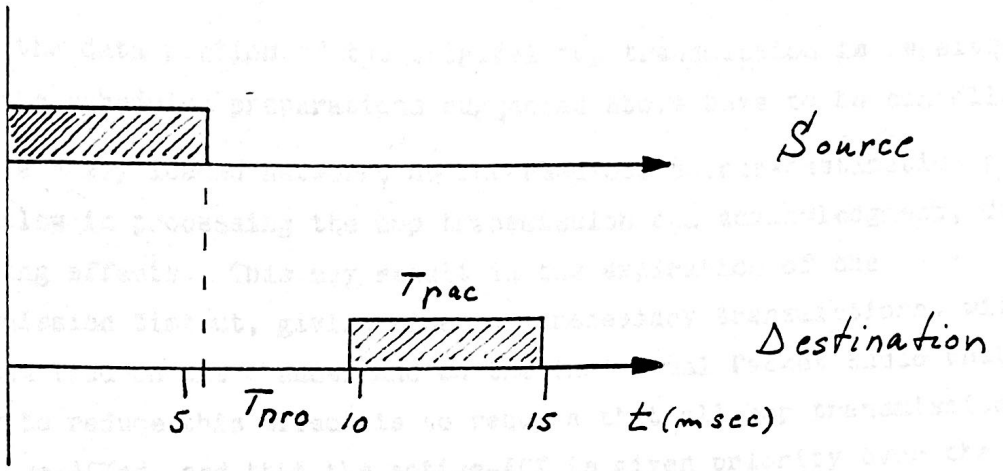

FIGURE 3 TRANSMISSION AND FORWARDING TIMES FOR AN INTERMEDIATE HOP

To clock out on the channel a full-size packet, takes 5.28 msec at 400 kb/s. When successfully received, the necessary time to process the header information and schedule the packet to be transmitted over the next hop, is measured to be about 4.2 msec. This is not much smaller than the transmission time, and is the main factor limiting the achievable throughput.

If the receiver can start to process the packet header information immediately after it is received, then while the rest of the packet is being received, it can make the necessary preparation - to forward the packet, or to schedule an active-ACK - in anticipation of a correctly received data portion. The forwarding or active-ACKing can then proceed shortly after the previous hop transmission has been fully received and verified, and thereby greatly improve the performance.

In order to perform the acknowledgment correctly, only the header part of the forwarded packet is needed at the source unit. The rest of

the packet may be in error, since it does not affect the acknowledgment of the previous hop. It is therefore necessary that the packet header is guarded by a separate checksum, if this scheme is going to be efficient.

If the data portion of the original hop transmission is received in error, the scheduled preparations suggested above have to be cancelled.

In a heavy loaded network, an intermediate source-destination pair may be slow in processing the hop transmission and acknowledgment, due to queuing effects. This may result in the expiration of the retransmission timeout, giving rise to unnecessary transmissions, with increased load on the channel and on the individual Packet Radio units. One way to reduce this effect is to require that all hop transmissions are active-ACKed, and that the active-ACK is given priority over the normal traffic. When a packet is successfully received, the receiver immediately respond by transmitting an acknowledgment. Further processing of the packet is postponed till the other tasks are finished.

The two methods can be combined. The first one is used when the output queue is empty. For a non-empty queue, an active-ACK is placed at the top of the queue, and if the packet has to be forwarded, it is placed at the end of it.

In the current CAP protocol, the interval trailing a packet transmission can be utilized to send packets by other units, thus multiplexing different data streams on the channel. The proposed method, to reserve this time interval for the acknowledgment, would improve the maximum throughput experienced by a single stream. It may therefore be a tradeoff between these two possible ways of utilizing the trailing time interval. But it is believed that the improved acknowledgment efficiency of the proposed method, counterweights the possible improvement in channel utilization by the current method, due to the reduction in retransmissions.

V PROPOSAL FOR ACTIVITY SIGNALLING AND IMPROVED ACKNOWLEDGMENT SCHEME

A. Summary of the Requirements

The requirements listed below are results of the previous discussions:

- The transmissions from any neighbour of a specific Packet Radio unit can be detected and successfully received by this unit, when no contention is experienced (symmetry requirement).
- The Busy Tone has the same range as normal transmitted packets.
- The Busy Tone is orthogonal with the data packets, in frequency or time.
- The Busy Tone is detected in the same way as packet transmissions.
- The Busy Tone contains no information.
- The packet header should be covered by a separate checksum.
- The receiving side should start processing the header information immediately after it is received, simultaneously with the reception of the data portion of the packet.

The first requirement is probably not obvious. It requires that any unit that can be heard by a particular radio unit, also should be inhibited by this unit when it is transmitting. If not, the carrier sense method would be less effective.

B. Scenarios for Transmitting and Receiving Sides for Case 2

Sending side:

- No carrier or Busy Tone detected, send packet immediately
- Carrier, or Busy Tone, or both, detected, do not send packet, reschedule for later transmission
- Channel becomes empty; i.e., ending of carrier or Busy Tone, wait an interval T and a random interval t, send the packet if the channel is empty.

Receiving side:

- Bit sync acquired, i.e., packet being received, start processing the header info. when received

```
If   received packet is destined for me
then turn on Busy Tone
else
     if   alternate routed packet and I can help
     then turn on Busy Tone
     fi;
fi;

if   packet ACKs a previous transmission from me
then update my tables
fi;

if   packet should be forwarded
then
     if   output queue is empty
     then place packet in the queue
     else place ACK at the top of the queue
          place packet at end of the queue
     fi;
else place ACK at the top of the queue
fi;

if   end of packet
then turn off Busy Tone
     if   packet is OK
     then proceed as prepared
     else cancel the preparations
fi;
```

VI CONCLUSIONS

The current note has discussed a method to extend the carrier sense technique by means of signalling of ongoing transmission and reception activities, in order to improve network performance under high loading conditions. In addition, various approaches to improve the efficiency of the present acknowledgment scheme are given. It has not yet been possible to estimate quantitatively the improvements in performance for the proposed methods, since this requires careful studies based on theoretical models and/or simulations. Such studies are currently under way.

In a military environment it is important that the Busy Tone have the same jam-resistance as the data channel, in order to prevent or reduce intentional blocking of the network. This requirement may not be as critical in commercial applications. The implementation of the concept is not achieved without penalties, and must be traded against the increase in performance. Principally there are two ways to implement the Busy Tone scheme, one is orthogonal in frequency and requires an extra set of transmitter/receiver pair in each Packet Radio unit, while the other is orthogonal in time and can therefore be multiplexed with the data. The complexities and consequences of the two methods are being investigated, in order to make the necessary tradeoffs.

REFERENCES

1. Kahn R., S. A. Gronemeyer, J. Burchfiel, and R. C. Kunzelman, "Advances in Packet Radio Technology," Proc. IEEE, Vol. 66, No. 11, (1978).

2. Kleinrock, L. and F. Tobagi, "Packet Switching in Radio Channels: Part I - Carrier Sense Multiple Access Modes and Their Throughput-Delay Characteristics," IEEE Trans. Commun., pp. 1400-1416, (December 1975).

3. Tobagi, F. and L. Kleinrock, "Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple-Access and Busy-Tone Solution," IEEE Trans. Commun., pp. 1417-1433 (1975).

4. Tobagi F., S. E. Lieberson, and L. Kleinrock, "On Measurement Facilities in Packet Radio Systems," NCC 1976, pp. 589-596.